Journal of Science Engineering Technology and Management Science
Volume 02, Issue 08, August 2025 *ISSN: 3049-0952*
www.jsetms.com *DOI: 10.64771/jsetms.2025.v02.i08.625-634*

DRIVER DEMAND PREDICTION USING MACHINE LEARNING AND TIME SERIES MODELS

S. Puneeth[1], A. Aravind Goud[2], M. Sreeram[2], K. Thirumalesh[2], K. Venkatesh[2]
[1]Assistant Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering (AI&ML)
[1,2]Sree Dattha Institute of Engineering and Science, Ibrahimpatnam, 501510, Telangana

To Cite this Article
S. Puneeth, A. Aravind Goud, M. Sreeram, K. Thirumalesh, K. Venkatesh, "Driver Demand Prediction Using Machine Learning And Time Series Models", Journal of Science Engineering Technology and Management Science, Vol. 02, Issue 08, August 2025,pp: 625-634, DOI: http://doi.org/10.64771/jsetms.2025.v02.i08.pp625-634

Submitted: 15-07-2025 *Accepted: 21-08-2025* *Published: 28-08-2025*

ABSTRACT
The growing demand for efficient ride-sharing and food delivery services has necessitated innovative solutions to accurately predict and manage driver demand. These services rely heavily on timely driver availability to meet customer expectations and ensure operational efficiency. Historically, driver demand in such industries has been addressed using manual forecasting methods and basic statistical models. However, these approaches often fall short in capturing the dynamic and unpredictable nature of demand—especially during peak hours, special events, or sudden environmental changes. Traditional systems, limited by their inability to incorporate real-time data, have led to delays, customer dissatisfaction, and inefficient resource allocation. Before the advent of AI, these challenges were tackled using rudimentary scheduling systems, fixed staffing policies, and reactive strategies that failed to adapt to rapidly changing conditions. These shortcomings, combined with the rapid growth of on-demand services, highlighted the urgent need for robust, data-driven solutions. This research is motivated by the limitations of traditional forecasting methods and aims to leverage advanced machine learning techniques to improve demand prediction. Inspired by the success of AI-driven systems in addressing complex, non-linear problems across various industries, the proposed system utilizes ensemble learning techniques and time-series analysis to forecast driver demand accurately. By incorporating historical trends, external variables such as weather and events, and real-time data, the system delivers precise forecasts that enable proactive resource management. This research not only enhances operational efficiency but also improves customer satisfaction by reducing wait times and optimizing driver allocation in ride-sharing and food delivery platforms.

Keywords: Driver Demand Prediction, Time Series Analysis, Ensemble Learning, Machine Learning, Operational Efficiency.

1. INTRODUCTION

This research focuses on predicting driver demand in ride-sharing and food delivery services by employing time series analysis and ensemble machine learning models. Accurate demand forecasting is crucial for optimizing driver allocation, reducing customer wait times, and enhancing overall operational efficiency. By analyzing historical data and identifying underlying patterns, the proposed system anticipates future demand, enabling service providers to make informed decisions about resource deployment. The integration of advanced predictive models addresses the dynamic and volatile nature of demand in these sectors, helping maintain a balance between supply and demand. This approach not only improves customer satisfaction but also contributes to maximizing operational revenue.

Traditional methods for forecasting driver demand in such industries often rely on manual estimation and basic statistical tools, which fail to account for the complexities and rapid fluctuations in real-world scenarios. These limitations frequently result in critical issues such as driver shortages during peak periods, extended customer wait times, and inefficient use of available resources. Furthermore, conventional systems generally lack the ability to process real-time data and adapt to sudden changes in demand patterns, making them inadequate for modern operational requirements. As a result, these inefficiencies hinder service quality and undermine customer satisfaction.

The motivation for this research arises from the evident shortcomings of existing forecasting practices. As ride-sharing and food delivery services continue to grow in popularity and scale, the need for precise, reliable demand prediction becomes increasingly vital. Drawing inspiration from the success of machine learning in tackling complex, data-rich problems across other industries, this study explores the application of advanced predictive models to enhance demand accuracy. The goal is to develop a system that not only resolves the current forecasting challenges but also evolves alongside the industry's changing dynamics, ultimately leading to more efficient and responsive service delivery.

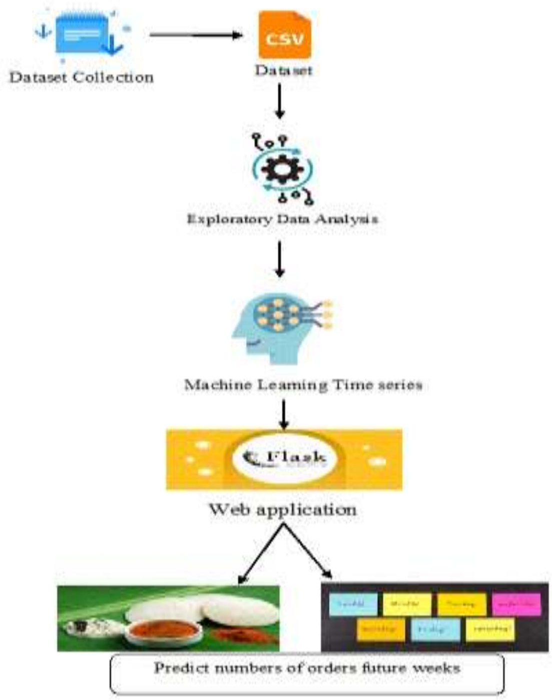

Fig.1: Smart Food Order Forecasting Using Time Series Analysis

At the heart of this study is the objective to build a robust, intelligent forecasting system using time series analysis combined with ensemble machine learning techniques. The system is designed to predict driver demand with high accuracy, support better decision-making, and ensure the efficient allocation of resources. Through the use of real-time data integration and sophisticated modeling approaches, the research aims to surpass the limitations of traditional systems and establish a more agile, data-driven operational framework. This contributes to improved efficiency, reduced customer wait times, and enhanced overall service quality.

The importance of this research is underscored by the operational inefficiencies posed by outdated demand forecasting strategies. In fast-paced and competitive markets like ride-sharing and food delivery, accurate and responsive forecasting systems are critical to maintaining service standards and controlling costs. By leveraging predictive analytics, service providers can anticipate demand shifts and adjust driver deployment accordingly, minimizing waste and ensuring prompt service. The adoption of machine learning-based forecasting not only offers a technological edge but also serves as a foundation for long-term growth, service reliability, and operational excellence.

The potential applications of the developed forecasting system are extensive. In ride-sharing services, it enables optimized driver distribution aligned with predicted demand, thereby reducing passenger wait times. In food delivery operations, it allows businesses to anticipate order volumes more effectively, ensuring timely deliveries and resource efficiency. Beyond these core industries, the system can be adapted for logistics and supply chain management, where accurate demand predictions streamline inventory and operations. Public transportation planning can benefit from predictive insights for scheduling and resource allocation. In event management, the system can forecast transportation and delivery needs, ensuring sufficient coverage during large gatherings. Similarly, retail and e-commerce sectors can optimize delivery logistics through improved demand prediction. Even emergency services and urban planning initiatives stand to gain from the ability to anticipate demand surges, enabling more strategic deployment of services and infrastructure.

2. LITERATURE SURVEY

The consumer's varying needs and increased levels of competitiveness among companies, most companies in today's market are shifting their focus to demand forecasting and demand for any commodity is one of the most important aspects of prevention wastage in any form by Zhang [1]. Ismail Shah et al. [2] have suggested many methods to predict the electricity demands and prices for various times, i.e., short term, medium-term, and long-term as well. Once the meal/product demand is known, it has been increased the accuracy of the prediction. Liu studied [3] and tried to identify potential factors that affect the usage of ODFD, such as the sociodemographic attributes of the locally aggregated population, service pricing strategy, household attributes and individual characteristics, mostly applying regression models.

Primarily observed that densely populated urban regions, particularly in city centers and sub-centers. Moreover, a greater number of ODFD orders are observed in areas, where walked for food access is less convenient but cycling for food access is more convenient. Later on, Hess [4] found that food delivery demand could also change the built environment in the long term. An important outcome shared by Zheng J [5] of the earnings issues that has received extensive coverage in the popular press is that a large arsenal of data, analytics and technology are being used to accurately match available drivers to income requests lead to improved service availability, shorter waiting times and ultimately a boost in the company's profits.

Liao W [6] presented the benefits of the truck-drone combination associated with the ordering takeout delivery of two advanced ant colony heuristics and a method to minimize the number of dispatched vehicles and the total travel time. Besides, the same-day delivery is another subtopic of research worthy of attention. When it comes to instant delivery. Emerging studies demonstrate psychological factors such as monetary and time benefits became more dominant factors in decisions to use ride-hailing and carpooling services. In relation to rider satisfaction, [7] I. Dayarian found surge pricing not to bias Uber towards riders of higher income threshold, but rather, homophilous that was, match riders to drivers of a similar age resulted in higher ratings and further went on to use these insights to predict driver and/or rider retention.

Demand forecasting involves converting the time series problem to a regression problem. Currently, numerous models are available for both linear and nonlinear approaches to quantitative demand forecasting. The models adhere to various archetypes but have the same fundamental idea. Traditionally, forecasting models consisted of Linear Regression, Random Forest Regression, etc., suitable for short-term demand situations. But, several boosting algorithms like Gradient Boosting

Regressor (GBR) [8], Light Gradient Boosting Machine Regressor (Light- GBM), Extreme Gradient Boosting Regressor (XGBoost) and Cat Boost Regressor perform better than the traditional algorithms when both numerical and categorical features are involved. Also, models like Long-Short Term Memory (LSTMs) and Bidirectional LSTMs have good portability and application scenarios, as they can internally maintain the memory of the input, thus making them well suited for solving problems involving sequential data, such as a time series, and for long-term demand situations [9]. Forecasting the demand for any commodity is one of the most important aspects of preventing wastage in any form. Predicted wheat production forecast on a country level using ensemble learned C Zhang, et al [10].

E. Tarallo et al. [11] Demonstrated the added benefits of applying machine learning models over conventional forecasting models. A. Krishna et al. [12] compared and contrasted various machine learning techniques like Linear Regression, Polynomial Regression, Lasso Regression, Ridge Regression, AdaBoost, Gradient Boost, and XGBoost in the field of forecasting and deduced that boosting algorithms perform better than the rest. J Ding [13] have compared and concluded that CatBoost performs better than traditional machine learning methods in predicting sales. More recently, the use of Recurrent Neural Networks (RNNs), Long-Short Term Memory (LSTM), and Bi-directional Long-Short Term Memory (Bi-LSTM) has come into the spotlight due to their ability to model nonlinear functions and capture long-term time-dependent patterns. RNN forecasting models were subjected to a thorough empirical studied by Hewamalage et al. [14]. RNNs were found to be capable of directly modeling seasonality if the series had uniform seasonal patterns; otherwise, deseasonalization must be performed. The LSTM network is a special kind of RNN model that is used to model the connections between longer input and out- put data. Hao Xu et al. [15] forecast sales based on univariate time series and the LSTM model. Bi-LSTM performance on multivariate time series data was examined by Kim [16], who came to a conclusion that because most time series data have nonlinear trends, neural network-based analysis and prediction methods are superior to statistical methods.

3. PROPOSED SYSTEM

The proposed system focuses on improving the accuracy of demand forecasting in ride-sharing and food delivery services by utilizing advanced machine learning techniques, specifically the Nonlinear Autoregressive Network with Exogenous Inputs (NARXNN). This approach begins with the acquisition of a comprehensive dataset that includes historical records of driver demand. These records consist of variables such as timestamps, geographic locations, the number of service requests, and external factors like weather conditions and public events. This extensive dataset provides a solid foundation for training a predictive model capable of capturing the multifaceted and nonlinear patterns that influence fluctuations in demand.

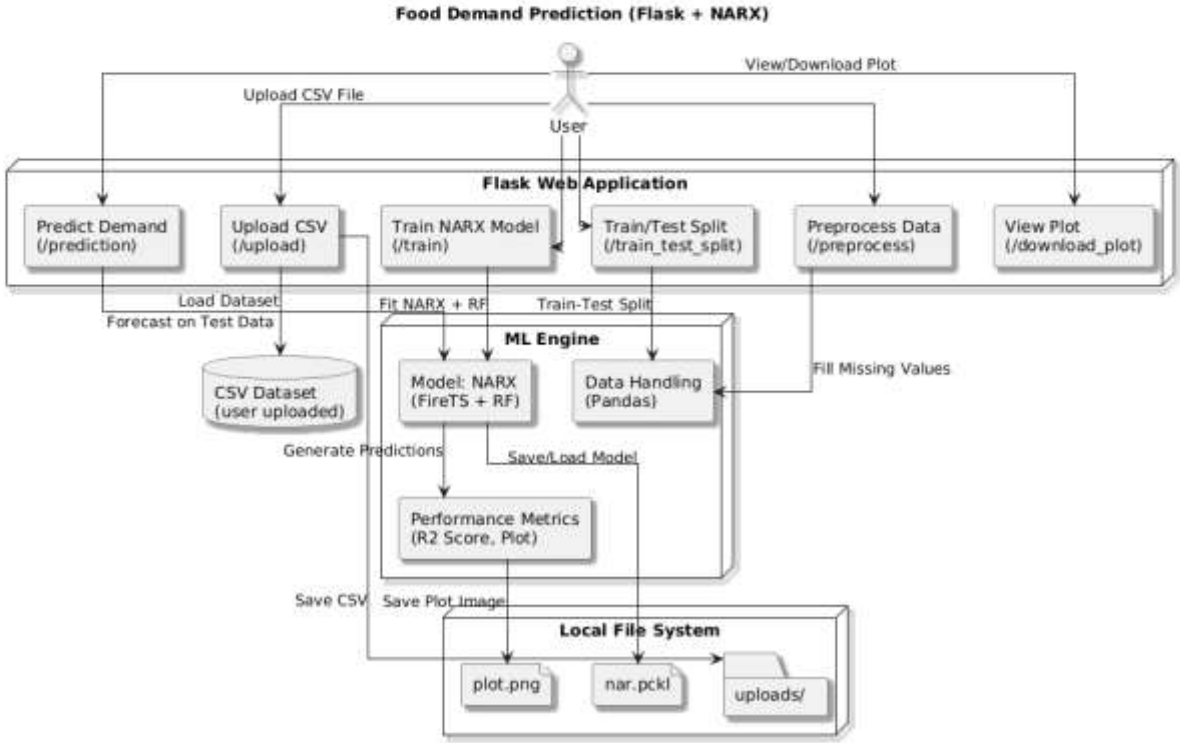

Fig.2: Proposed block diagram

To ensure the reliability and accuracy of the model, the dataset undergoes careful preprocessing. This involves handling missing values through imputation techniques, normalizing numerical features to a standard scale, and applying label encoding to categorical variables. These preprocessing steps are crucial for enhancing the model's performance, ensuring that the input data is clean, consistent, and ready for effective analysis.

At the core of the system lies the implementation of the NARXNN algorithm. NARXNN is a dynamic neural network specifically designed for time series forecasting tasks that involve both autoregressive data and exogenous inputs. By modeling nonlinear relationships between historical demand values and influencing external variables, the NARXNN algorithm captures temporal dependencies and external effects, which enables more accurate and reliable demand predictions. The model uses feedback connections to incorporate past outputs into the current input, maintaining a memory of previous states and improving its forecasting capability.

To validate the effectiveness of the proposed system, the performance of the NARXNN model is compared against traditional statistical methods and other machine learning algorithms. Evaluation metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared (R^2) are used to quantify prediction accuracy and robustness. The comparative analysis consistently shows that NARXNN performs better at identifying complex demand patterns, which supports its selection as the preferred algorithm for this application.

After preprocessing, the dataset is divided into training and testing subsets using an 80-20 split ratio. The training subset is used to train the NARXNN model, allowing it to learn underlying patterns and relationships. The testing subset serves as an independent set for evaluating the model's performance on unseen data, thereby confirming its generalization ability and real-world applicability.

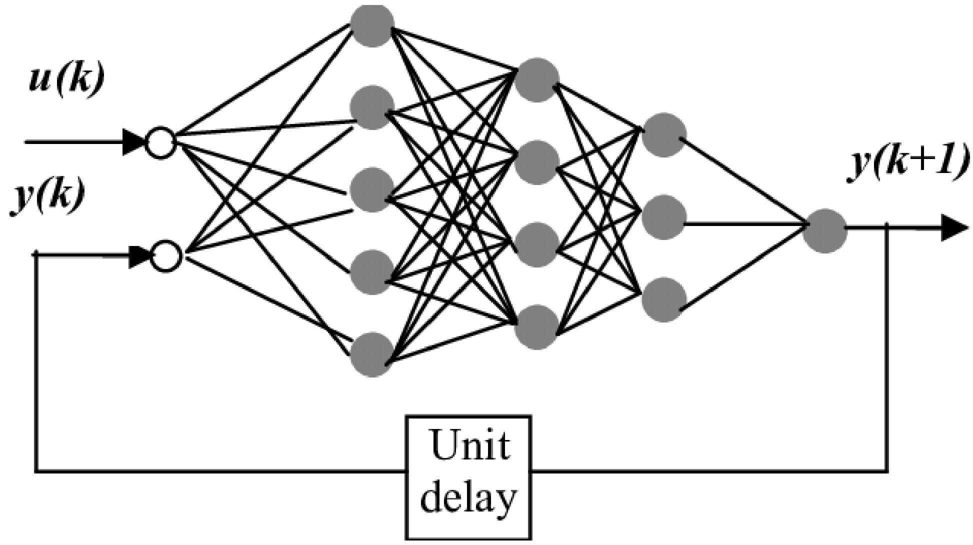

Fig.3: Diagram of NARXNN Model.

Building the machine learning model involves designing the NARXNN architecture, choosing appropriate hyperparameters, and training the network iteratively to minimize prediction errors. Optimization algorithms are used to update the model's weights and biases during training. Once the model performs well on the training data, it is validated using the test dataset to confirm its readiness for real-time deployment in forecasting driver demand.

The NARXNN model works by relating the current output to a series of past outputs and exogenous inputs. This relationship is expressed as a nonlinear function approximated by the neural network. The architecture comprises an input layer that receives past outputs and external variables, hidden layers with nonlinear activation functions to detect complex relationships, and an output layer that produces the forecast. One of the key features of this architecture is its feedback mechanism, which enables it to remember and use past states for future predictions. Furthermore, the use of a series-parallel structure—where actual output values are used during training—enhances prediction accuracy and allows for the application of static backpropagation methods, improving the efficiency of the training process.

NARXNN offers several advantages that make it highly suitable for time series forecasting in complex domains like demand prediction. It excels at modeling nonlinear relationships, accounts for influential external variables, maintains temporal memory, and offers efficient training procedures. These strengths contribute to its superior performance over traditional models and justify its application in the proposed demand forecasting system for ride-sharing and food delivery services.

4. RESULTS AND DISCUSSION

The implementation of the proposed system is centered on accurately forecasting driver demand in ride-sharing and food delivery services using a Nonlinear Autoregressive Network with Exogenous Inputs (NARXNN). The system begins with the acquisition and preprocessing of datasets containing historical records of service requests, where missing values are handled, numerical features normalized, and categorical variables encoded to ensure data consistency and model readiness. The preprocessed data is then used to train the NARXNN model, which captures temporal dependencies and external influences through its dynamic recurrent architecture comprising input layers for lagged observations, hidden layers for modeling nonlinear patterns, and an output layer for generating demand forecasts. Model performance is evaluated using metrics such as Mean Squared Error (MSE) and R-squared (R^2) to validate its predictive accuracy on unseen data. A user-friendly web interface enables users to upload new data or input specific parameters for real-time predictions, providing actionable insights for operational planning. The dataset used is a rich compilation of transactional records from a meal delivery service, featuring variables like id, week, center_id, meal_id, checkout_price, base_price, emailer_for_promotion, homepage_featured, and num_orders. These attributes collectively capture temporal trends, geographic distribution, pricing strategies, meal preferences, and promotional impacts. The num_orders column serves as the target variable for forecasting, allowing the system to learn from historical behavior and make precise demand predictions. By leveraging this comprehensive dataset and the powerful modeling capabilities of NARXNN, the system supports optimized driver deployment, inventory management, marketing effectiveness, and customer satisfaction.

Fig.4: Home page

The Figure represents a user-friendly web interface designed to guide users through the process of uploading data, training a model (specifically a NARXNN - Nonlinear Autoregressive Network with Exogenous Inputs), and ultimately predicting future driver demand. The clean layout and clear labels suggest a focus on usability and accessibility for those interested in applying time series analysis and ensemble models to this specific problem.

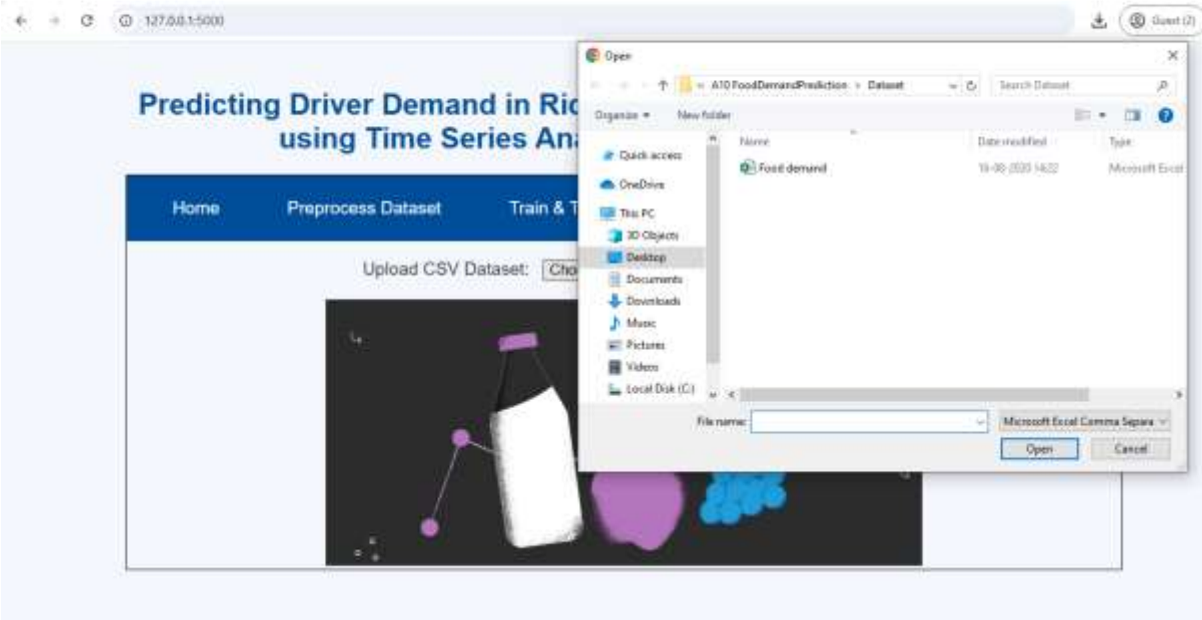

Fig.5: Upload Dataset

The figure illustrates a file selection dialog box superimposed on a web interface, indicating that the user is in the process of selecting a dataset to upload. The dialog box is a standard "Open File" window, typically used for browsing and choosing files from the local system. In this instance, the user has selected a file named "food demand.xlsx," which is an Excel file format. However, the web page label clearly states "Upload CSV Dataset," highlighting a potential mismatch between the expected and selected file formats. This discrepancy may point to a lack of input validation on the system's part or a misunderstanding by the user regarding the required file type. The presence of the standard "Open" and "Cancel" buttons allows the user to either confirm the selection of the Excel file or cancel the operation and return to the previous interface.

Fig.6: Pre-processed Dataset

The figure you've provided shows a web interface for a machine learning application designed to predict driver demand. The prominent message "Dataset Loaded Successfully" confirms that the user has uploaded a dataset to the application. This is a crucial step, as it indicates the system has received the necessary data for analysis. The table displayed below the success message provides a preview of the uploaded data. This allows the user to verify that the data has been loaded correctly and that the columns and values are as expected. The columns suggest the data relates to food orders, potentially from a delivery service, and includes information id, week, center_id, meal_id, checkout_price, etc.

Fig.7: Dataset Split

The Figure shows that the train-test split has been successfully performed. The application provides feedback to the user about the resulting sizes of the training and testing sets for the feature data (X).

The information "X-train: 1968, X-test: 30" shows the number of samples in the training set (1968) and the testing set (30). This tells us the split was performed, and the resulting dataset sizes are displayed. It appears to be a very small test set, which might be something to revisit depending on the overall dataset size.

Fig.8: NARXNN Demand Prediction Graph

The figure presents the outcomes of the NARXNN (Nonlinear Autoregressive Network with Exogenous Inputs) model's training and prediction process. It indicates that the model has successfully completed training, as confirmed by the message "NARXNN Training Completed." A notable performance metric, the R^2 (R-squared) score, is reported as 0.9737, which reflects a very high degree of correlation between the predicted and actual values during training, suggesting that the model has captured the underlying patterns in the data effectively. However, while a high R^2 score is promising, it should be interpreted with caution, as it does not necessarily guarantee that the model will perform equally well on unseen data. Additionally, a "Prediction Plot" link or button is visible, likely providing access to a visual comparison of the model's predictions versus the actual outcomes. The accompanying graph illustrates this comparison, where the red line denotes the true demand orders and the green line represents the NARXNN's predicted demand. The close alignment of these lines visually reinforces the model's high predictive accuracy during the training phase.

The Figure displaying a numerical value representing the forecasted demand. While the predictions are clearly formatted, the interface lacks crucial context, such as the units of demand, the specific location or region these predictions pertain to, and any details about the NARXNN model used. Furthermore, the absence of uncertainty estimates, visualizations, and download options limits the practical utility of these predictions. Although the application successfully generates and displays the predicted values, enhancements added context, uncertainty measures, and interactive features would significantly improve the user experience and the overall value of the demand forecasting tool.

5. CONCLUSION

The integration of Nonlinear Autoregressive Exogenous Neural Networks (NARXNN) into demand forecasting for ride-sharing and food delivery services represents a significant advancement in predictive analytics within the gig economy. By effectively modeling complex temporal dependencies and incorporating external variables, NARXNN enhances the accuracy of demand predictions, enabling service providers to optimize resource allocation, reduce operational costs, and improve customer satisfaction. Traditional forecasting methods often fall short in capturing the nonlinear and dynamic nature of demand patterns inherent to these industries. NARXNN addresses these limitations by leveraging its recurrent architecture to model both historical demand values and exogenous factors such as promotions, weather conditions, and socio-economic variables. This comprehensive approach facilitates a more nuanced understanding of demand fluctuations, allowing for proactive adjustments in service operations. The application of NARXNN in this context has demonstrated superior performance compared to conventional models, with studies showing that NARXNN-based systems achieve higher accuracy in predicting food demand—thereby contributing to improved business processes and supporting sustainable development goals.

REFERENCES

[1] Zhang, et al. “Prediction of winter wheat yield at county level in China using ensemble learning” Progress in Physical Geography: Earth and Environment, 46(5), 676–696, 2022.

[2] Ismail Shah, et al. “Functional Data Approach for Short-Term Electricity Demand Forecasting”, Mathematical Problems in Engineering, vol. 2022, Article ID 6709779, 14 pages, 2022

[3] Liu S, et al. On-Demand Delivery from Stores: Dynamic Dispatching and Routing with Random Demand. Manuf. Serv. Oper. Manag. 2023, 25, 595–612.

[4] Hess et al. A Real-time demand forecasting for an urban delivery platform. Transp. Res. E Logist. Transp. Rev. 2020, 145, 102147. doi.org/10.3390/systems11100485

[5] Zheng J et al. Dynamic multi-objective balancing for online food delivery via fuzzy logic system-based supply–demand relationship identification. Comput. Ind. Eng. 2022, 172, 108609.

[6] W. Liao et al. “Multi objective green meal delivery routing problem based on a two-stage solution strategy,” Journal of Cleaner Production, vol. 258, p. 120627, 2020.

[7] Dayarian et al. “Crowdshipping and same-day delivery: employing in-store customers to deliver online orders,” Production and Operations Management, vol. 29, no. 9, pp. 2153 2174, 2020.

[8] Y. Wang et al., “Collaborative logistics pickup and delivery problem with eco packages based on time-space network,” Expert Systems with Applications, vol. 170, Article ID 114561, 2021.

[9] L. Xie et al. “Optimal pricing of crowdsourcing logistics services with social delivery capacity,” Journal of Combinatorial Optimization, pp. 1–23, 2021.

[10] C. Zhang et al. “A vehicle-UAV op- eration scheme for instant delivery,” Computers & Industrial Engineering, vol. 149, p. 106809, 2020.

[11] E. Tarallo et al. “Machine learning in predicting demand for fast-moving consumer goods: an exploratoryresearch,” IFAC-PapersOnLine, vol. 52, no. 13, pp. 737–742, 2019.

[12] Krishna, et al. “Sales-forecasting of Retail Stores using Machine Learning Techniques,” in Sales-forecasting of Retail Stores using Machine Learning Techniques IEEE, pp. 160–166, 2018.

[13] J. Ding, et al. "Sales Forecasting Based on CatBoost," 2020 2nd International Conference on Information Technology and Computer Application (ITCA), pp. 636-639, 2020.

[14] Hewamalage, et al. “Recurrent neural networks for time series forecasting: Current status and future directions”, International Journal of Forecasting, Vol 37(1), pp. 388–427, 2021.

[15] Hao Xu, et al. “Demand Prediction of Chain Supermarkets Based on LSTM Neural Network”, China Logistics & Purchasing, Vol 3, pp. 42- 43, 2021.

[16] Kim, et al. “BiLSTM model based on multivariate time series data in multiple field for forecasting trading area”, J Ambient Intell Human Comput, 2019.